

# STATISTICAL REPORTS

*Ecology*, 87(4), 2006, pp. 835–841  
© 2006 by the Ecological Society of America

## GENERALIZED SITE OCCUPANCY MODELS ALLOWING FOR FALSE POSITIVE AND FALSE NEGATIVE ERRORS

J. ANDREW ROYLE<sup>1</sup> AND WILLIAM A. LINK

*U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA*

**Abstract.** Site occupancy models have been developed that allow for imperfect species detection or “false negative” observations. Such models have become widely adopted in surveys of many taxa. The most fundamental assumption underlying these models is that “false positive” errors are not possible. That is, one cannot detect a species where it does not occur. However, such errors are possible in many sampling situations for a number of reasons, and even low false positive error rates can induce extreme bias in estimates of site occupancy when they are not accounted for. In this paper, we develop a model for site occupancy that allows for both false negative and false positive error rates. This model can be represented as a two-component finite mixture model and can be easily fitted using freely available software. We provide an analysis of avian survey data using the proposed model and present results of a brief simulation study evaluating the performance of the maximum-likelihood estimator and the naive estimator in the presence of false positive errors.

**Key words:** *finite mixture; heterogeneous detection probability; latent class model; multinomial misclassification; multinomial mixtures; nondetection bias; occurrence probability; proportion of area occupied.*

### INTRODUCTION

The problem of estimating occurrence probability, proportion of area occupied (PAO), or “site occupancy,” of a species subject to imperfect detection is a problem of some interest in many animal sampling problems (Bayley and Peterson 2001, MacKenzie et al. 2002, Nichols and Karanth 2002). As incremental methodological extensions have been developed (e.g., Royle and Nichols 2003, MacKenzie et al. 2003, 2005, Tyre et al. 2003, Gu and Swihart 2004, MacKenzie and Bailey 2004, Royle 2004a, b), the site occupancy framework is becoming widely adopted in survey and monitoring activities in many settings. For example, the Amphibian Research and Monitoring Initiative (ARMI; Hall and Langtimm 2001) has identified site occupancy as the primary focus of current efforts as indicated by the following statement from the ARMI web page (*available online*):<sup>2</sup>

*The most promising national variable to date is one based on species presence. Documenting shifts in species presence through time will provide important data for assessing changes in amphibian status. The*

*“proportion of area occupied” (PAO) by an amphibian species has been identified by ARMI as the only metric that so far meets the Program criteria for being nationally interpretable and regionally adaptable.*

(See also Swihart et al. [2003], Bailey et al. [2004], Weir et al. [2005], Ball et al. [2005], and Stanley and Royle [2005] for other applications, including to other taxa). The rapid and widespread adoption of site occupancy as a metric of animal population status is due, at least in part, to the ease of establishing surveys based on presence/absence data, that such models facilitate an explicit accounting for detectability of the species in question, and the extensibility of the site occupancy modeling framework.

The framework laid out by MacKenzie et al. (2002) and others addresses one fundamental concern in animal sampling problems: the problem of false negatives in survey data, or the failure to detect a species where in fact the species is present. As an example, suppose a sample unit (or “site”) is sampled multiple times, yielding a record of putative absence such as, for  $T = 3$  visits, (0, 0, 0). There are two mutually exclusive possibilities to explain this event. The first is that the species is actually absent from the site, and thus the survey result was accurate. The second is that the species was in fact present at the site but went undetected (a

Manuscript received 13 June 2005; revised 28 October 2005; accepted 31 October 2005. Corresponding Editor: F. He.

<sup>1</sup> E-mail: aroyle@usgs.gov

<sup>2</sup> <http://armi.usgs.gov/monitoring.asp>

“false absence”). This is fairly typical in surveys of wildlife, where, of course, vast literature has developed around the issue of modeling detectability wildlife sampling (e.g., see Williams et al. 2002). The development of a formal framework for modeling and inference based on data subject to false negative errors was the focus of the work by Bayley and Peterson (2001), MacKenzie et al. (2002, 2003), Royle and Nichols (2003), Tyre et al. (2003), and others. These approaches have been predicated on the assertion that false positives were impossible, i.e., that the species will not be detected where it does not occur. While this assumption might appear to be obviously true, the fact is that misidentification can occur in field settings. If false positives do occur, it is absolutely critical that they be accommodated in the model. Otherwise, under a simple binomial sampling scheme, apparent occupancy will tend to 1.0 as the number of visits to sites increases. Indeed, even for very low rates of false positive errors, the bias in apparent occupancy rates, or those estimated under a model that does not permit false positives, will be extreme (see *Simulation study*). A second reason that we feel interest should be paid to false positives is that many surveys, in particular of birds and anurans, involve the simultaneous sampling of large numbers of species by volunteer observers with highly varying skill levels. This circumstance is ideal for the introduction of false positives by misidentification of species. We thus suggest that false positives are likely to be common in many survey situations where they have previously been disregarded.

In this paper, we generalize the site occupancy model described by MacKenzie et al. (2002) and others to the situation where both false negative and false positive observations are possible. In *A model for misclassification*, we present the model for false positive errors, noting that it can be represented as a finite mixture of binomial random variables. We note that the framework for modeling false positive errors is more general and does extend to situations wherein there are more than two possible states (e.g., Royle and Link 2005). This is briefly discussed below in *A model for misclassification:  $K > 2$  classes*.

#### A MODEL FOR MISCLASSIFICATION

Suppose that  $i = 1, 2, \dots, R$  units (sites) are classified according to some characteristic having  $K$  possible values. Denote the true class of unit  $i$  as  $z_i$  and the observed class of unit  $i$  for time  $t$  as  $y_{it}$ . Our focus here is on the case  $K = 2$ , although  $K > 2$  can be handled similarly. For  $K = 2$ , the possible classes are “occupied” ( $z = 1$ ) or “not occupied” ( $z = 0$ ). We suppose that  $y_{it}, t = 1, 2, \dots, T$  are independent and identically distributed, in which case, the sufficient statistic is a function of the site-specific totals  $y_i = \sum_t (y_{it})$ . We will focus on estimating the parameter  $\psi = \Pr(z_i = 1)$ , which is the site occupancy or proportion of area occupied (PAO) parameter considered by MacKenzie et al. (2002), or simply the probability of occurrence. As in MacKenzie

TABLE 1. Classification probabilities for site occupancy model with two possible values (i.e.,  $K = 2$ ).

$y_{it}$	$z$	
	0 (unoccupied)	1 (occupied)
0 (not detected)	$p_{00}$	$p_{01}$
1 (detected)	$p_{10}$	$p_{11}$

et al. (2002), we assume that occupancy statuses of sites are independent, and that the system is closed in the sense that occupancy status does not change during the period of sampling. The extension to open systems is straightforward (MacKenzie et al. 2003), and the independence assumption is conceptually helpful, but is not necessary in the sense that  $\psi$ , the MLE of  $\psi$ , is consistent even if occupancy statuses are not independent (e.g., if samples are too close together in space).

Fundamental to the problem considered here is that the observed class might not be equal to the true class for each unit. Consequently, we define the set of (mis)classification probabilities

$$p_{kl} = \Pr(y_{it} = k | z_i = l)$$

which sum to 1 for each value of  $z_i$ . These classification probabilities, arranged in tabular form for  $K = 2$ , are shown in Table 1, where each column constitutes the sampling distribution of  $y$  conditional on a particular state of  $z$ . Note that  $p_{00} = 1 - p_{10}$  and  $p_{01} = 1 - p_{11}$ . This model implies that, conditional on occupancy state (i.e., the value of  $z_i$ ), the site-specific counts  $y_i$  have a binomial distribution with a state-dependent detection probability parameter. That is, for an occupied site,  $y_i$  is binomial with parameter  $p_{11}$ , (“detection probability”) whereas, for an unoccupied site,  $y_i$  is binomial with parameter  $p_{10}$ . Thus,  $p_{10}$  is the probability of falsely detecting the species at an unoccupied site, i.e., the false positive rate parameter. Given data  $\mathbf{y} = \{y_i\}_{i=1}^R$  a likelihood for  $\mathbf{p} = (p_{11}, p_{10})$  and  $\psi$  can be written explicitly as

$$L(\mathbf{p}, \psi | \mathbf{y}) \propto \prod_{i=1}^R \left\{ \left[ p_{11}^{y_i} (1 - p_{11})^{T-y_i} \right] \psi + \left[ p_{10}^{y_i} (1 - p_{10})^{T-y_i} \right] (1 - \psi) \right\}. \quad (1)$$

The occupancy model described by MacKenzie et al. (2002) is a special case of the misclassification model, that being the model that arises under the constraint  $p_{10} = 0$ .

The right-hand side of Eq. 1 can be maximized numerically to obtain MLEs of the parameters  $p_{10}$ ,  $p_{11}$ , and  $\psi$ . Further, standard procedures for likelihood-based inference can be applied to this model. For example, estimated standard errors can be obtained from the Fisher information matrix evaluated at the MLEs (e.g., see Williams et al. 2002: appendix D). These can be used to construct confidence intervals for parameters of interest. Also, given a set of candidate

models, model selection can be carried out using Akaike's information criterion (AIC; e.g., see Burnham and Anderson 1998).

### *Representation as a finite mixture*

The explicit construction of the misclassification model given by Eq. 1, in which the data are binomial counts with state-specific detection probabilities, is commonly referred to as a finite mixture or latent class model. In the present case, the latent (i.e., not necessarily observable) class is species presence or absence from a site. Such models are common in many application areas as a mechanism for accommodating overdispersion in binomial counts. The finite mixture has been considered in similar contexts by Norris and Pollock (1996) and Royle (2005) (see also Link 2003). We note that the present application is somewhat more conventional because Eq. 1 is a regular binomial mixture with two classes whereas in the Norris and Pollock (1996) application, the index  $N$  is an unknown parameter (population size) to be estimated.

The main implication of this representation as a finite mixture is the existence of certain symmetries in the likelihood and the multimodality that this induces. That is, there is identical support for multiple, but distinct, sets of parameter values. It can be seen by inspection of Eq. 1 that  $L(p_{11}, p_{10}, \psi) = L(p_{10}, p_{11}, 1 - \psi)$ . Thus, for example, a detection rate of 70% at occupied sites and false detection rate of 40%, with 80% of sites occupied, has the same likelihood as 40% detection rate at occupied sites and 70% false detection rate, with 20% of sites occupied. A solution to this problem is to restrict the parameter space or, equivalently, to choose among equally well-supported alternatives. One sensible solution (to us) is to assert that  $p_{11} > p_{10}$ , i.e., that  $p$  is higher for occupied sites than the misclassification probability is for unoccupied sites. Assuming that  $p_{11} > p_{10}$  allows one to interpret the group with the higher binomial probability as the occupied group. This seems the most sensible solution because there is no context to group identification other than their values of  $p_{11}$  and  $p_{10}$ . That is, in the presence of misclassification, one does not really know definitively whether any site is occupied. The converse could also be asserted ( $p_{10} > p_{11}$ ) a priori, but our point is that there is no information in the data to associate meaning with the model parameters. Note that the conventional site occupancy model avoids this difficulty by asserting that  $p_{10} = 0$ . Finally, this likelihood symmetry implies also that, as  $p_{11}$  gets close to  $p_{10}$ , the MLE become unstable.

### *$K > 2$ classes*

The model for misclassification described previously can be extended to the situation where there are  $K > 0$  possible states. For example, Royle (2004b) and Royle and Link (2005) consider models for anuran calling survey data. In these surveys, calling intensity of breeding anurans is recorded into a multinomial index

taking on values 0 (no calling activity) to 3 (full chorus). In both papers, the multistate equivalent of "false positive" errors were ruled out a priori. Indeed, Royle (2004b) asserted, incorrectly, that this restriction was necessary in order to decompose variation due to detectability from that due to population influences. In fact, we can develop a misclassification model for this  $K = 4$  situation by mixing a multinomial sampling distribution over potential states of the latent state variable  $z \in \{0, 1, 2, 3\}$ .

Specifically, let  $y_i \in \{0, 1, 2, 3\}$  be a categorical observation of abundance class and suppose that  $\Pr(y_i = k|z) = \pi_{k,z}$ ;  $k = 0, 1, 2, 3$  are the multinomial cell probabilities with  $\pi_{3,z} = 1 - \sum_{k=0}^2 \pi_{k,z}$  and  $z$  is the true but unknown state with  $\psi_k = \Pr(z = k)$ ;  $k = 0, 1, 2, 3$ . Royle and Link (2005) developed a modeling and inference framework for this situation by imposing the constraint that  $\pi_{k,z} = 0$  for  $k > z$ . However, the unconstrained model is identifiable provided that  $T$  is sufficiently large. The main difficulty when  $K > 2$  is that the multimodality of the likelihood is considerably more complex. The structure of the multinomial mixture mixed likelihood for  $K > 2$  is currently under investigation.

### EXAMPLE

Here, we present a brief example using detection/nondetection data on several species of passerines from a North American BBS Route composed of 50 spatial samples ("stops"), sampled 11 times during approximately a one-month sample period. We consider data on five species: Blue Jay, Common Yellowthroat, Song Sparrow, Gray Catbird, and Ovenbird (BLJA, COYE, SOSP, GRCA, and OVEN, respectively). Some of these data were analyzed in other contexts, using other models, by Royle and Nichols (2003) and also Royle (2005).

For each species, we fit the unconstrained model having likelihood given by Eq. 1 and the reduced model with the constraint  $p_{10} = 0$ , i.e., the model proposed by MacKenzie et al. (2002). Parameter estimates, standard errors for  $\hat{\psi}$ , and AIC for both models and for each species are given in Table 2. The models were fit using the free software package R (R Development Core Team 2005); the data and programs are made available in the Supplement to this paper.

For BLJA and GRCA, the misclassification model is not favored (although for GRCA, the AIC is about the same for both models). For the remaining species, the misclassification model is strongly favored. For SOSP, OVEN, and COYE there are large differences in estimated occupancy between the two models. For example, for COYE, estimated occupancy decreases from 0.723 under the constrained model to 0.364 for the misclassification model, and the estimated misclassification rate is 0.101. For SOSP, the misclassification is only 0.019, yet the estimated occupancy changes from 0.521 (constrained model) to 0.411 for the misclassification model. Apparently, very small rates of misclassification

TABLE 2. Parameter estimates and AIC under the proposed model allowing for false positive errors and the conventional model (with constraint  $p_{10} = 0$ ) fitted to five species of avian presence/absence data.

Species	$n/R$	Misclassification model					Constrained model			
		$\psi$	SE	$p_{11}$	$p_{10}$	AIC	$\psi$	SE	$p$	AIC
BLJA	0.66	0.694	0.159	0.204	0.007	170.04	0.723	0.077	0.199	168.08
COYE	0.72	0.364	0.080	0.587	0.101	225.63	0.723	0.064	0.385	254.55
SOSP	0.52	0.411	0.072	0.504	0.019	201.20	0.521	0.071	0.419	219.08
GRCA	0.38	0.307	0.087	0.264	0.012	136.23	0.407	0.075	0.219	136.11
OVEN	0.84	0.425	0.072	0.830	0.165	259.41	0.840	0.052	0.533	379.75

Notes: Species abbreviations are BLJA, Blue Jay; COYE, Common Yellowthroat; SOSP, Song Sparrow; GRCA, Gray Catbird; and OVEN, Ovenbird. The  $n/R$  column reports the observed proportion of occupied sites in the sample,  $n = \sum_{i=1}^R I(y_i > 0)$ ; and  $\psi$  is the site occupancy parameter or proportion of area occupied.

can have large effects on the apparent occupancy. Moreover, estimated occupancy under the misclassification model can be less than the apparent occupancy rate (i.e., the number of occupied sites observed in the sample).

SIMULATION STUDY

We conducted a small simulation study with two objectives: (1) to illustrate the bias of the standard occupancy estimator when misclassification is ignored and (2) to evaluate bias and precision of estimates of occupancy under the misclassification model. For clarity, we focus on a concise set of design and parameter settings for evaluation (the interested reader may deploy the R code provided in the Supplement to conduct additional simulations). For the design, we fix  $T = 5$  for all simulations and consider  $R = 100, 200, 500$  spatial replicates. We suppose three levels of detection probability (0.4, 0.6, 0.8) and two levels of misclassification probabilities (0.05, 0.10). This yields 18 design

levels for a fixed value of  $\psi$ . Selected results from the simulation for  $\psi = 0.6$  are given in Table 3.

These results demonstrate that, even with relatively small misclassification rates, there is extreme bias in the conventional estimator of site occupancy. Second,  $\hat{\psi}$  is approximately unbiased over the range of conditions that summarized in Table 3. Third, in terms of precision,  $\hat{\psi}$  appears reasonable even for  $R = 100$  except in the case  $p = 0.40$  (the lowest value of  $p$  considered).

DISCUSSION

Interest in site occupancy models that account for imperfect detection has increased dramatically over the last several years as useful extensions of these models have been introduced. These models have been developed primarily to address the important animal sampling problem of non-detection, or the existence of false negative errors. That is, given that a species is present at a site, it may go undetected. The most important and fundamental assumption implicit in the use of these methods is that false positive observations

TABLE 3. Summary of MLE of  $\psi$  under the model allowing for misclassification (mean and sd of the sampling distribution are given) and under a model in which misclassification rate is assumed to be 0 (only the mean of the sampling distribution is given).

No. sites	Simulation settings				Misclassification model		Constrained model
	$T$	$\psi$	$p_{11}$	$p_{10}$	$\hat{\psi}$	SD	$\hat{\psi}$
500	5	0.6	0.4	0.10	0.600	0.112	0.818
500	5	0.6	0.4	0.05	0.599	0.069	0.723
500	5	0.6	0.6	0.10	0.600	0.038	0.780
500	5	0.6	0.6	0.05	0.599	0.029	0.698
500	5	0.6	0.8	0.10	0.598	0.025	0.767
500	5	0.6	0.8	0.05	0.600	0.021	0.690
200	5	0.6	0.4	0.10	0.581	0.167	0.823
200	5	0.6	0.4	0.05	0.600	0.100	0.723
200	5	0.6	0.6	0.10	0.600	0.058	0.782
200	5	0.6	0.6	0.05	0.598	0.046	0.698
200	5	0.6	0.8	0.10	0.597	0.038	0.767
200	5	0.6	0.8	0.05	0.599	0.036	0.691
100	5	0.6	0.4	0.10	0.562	0.214	0.820
100	5	0.6	0.4	0.05	0.584	0.139	0.724
100	5	0.6	0.6	0.10	0.601	0.083	0.782
100	5	0.6	0.6	0.05	0.602	0.067	0.702
100	5	0.6	0.8	0.10	0.596	0.053	0.764
100	5	0.6	0.8	0.05	0.603	0.050	0.694

Notes: Results are based on 500 Monte Carlo samples for all cases. Variables are:  $T$ , the number of site visits;  $\psi$ , site occupancy parameter or proportion of area occupied;  $p_{11}$ , detection probability;  $p_{10}$ , the probability of falsely detecting the species at an unoccupied site, i.e., the false positive rate parameter;  $\hat{\psi}$ , the maximum likelihood estimator of  $\psi$ .



(i.e., “false presence”) are not possible. In this paper we presented a straightforward extension of the basic site occupancy model that allows for both false positives and false negatives. The model developed here contains one additional parameter,  $p_{10}$ , the false positive detection probability. When  $p_{10} = 0$ , the model reduces to that proposed by MacKenzie et al. (2002).

Several important generalizations of the model are, in principle, straightforward. For example, the inclusion of covariates that influence both  $\psi$  and  $p_{11}$  or  $p_{10}$ , and the extension to “open” systems pose no technical difficulty following the basic likelihood framework laid out by MacKenzie et al. (2002, 2003). The model for false positives also generalizes to multinomial classification data where  $K > 2$ , although we note that the dimensionality of such models becomes unwieldy and the data requirements are likely to be extreme. We have neglected the development of these extensions here in order to provide a clear and concise description of the misclassification problem and a framework for modeling false positive errors.

We believe that surveys of many taxa are prone to false positives due to misidentification of species because observers are required to collect data on many species simultaneously, and often several sympatric species may have similar visual or aural queues. For example, in anuran surveys conducted in Maryland, two species of grey treefrog occur (*Hyla versicolor* and *Hyla chrysoscelis*) that have similar vocalizations; the wood frog (*Rana sylvatica*), southern leopard frog (*Rana sphenoccephala*), and northern leopard frog (*Rana pipiens*) have similar calls, and Fowler’s toad (*Bufo fowleri*) and eastern narrowmouth toad (*Gastrophryne carolinensis*) can easily be confused. The problem of multispecies surveys in which species might be easily confused is exacerbated by having volunteer observers with highly variable skill levels. These same arguments can be made to support the plausibility of false positive errors in many avian survey programs. Unfortunately, the existence of even small rates of false positive errors can lead to severe bias in the estimator of site occupancy suggested by MacKenzie et al. (2002). Indeed, the probability of falsely concluding that a site is occupied in  $T$  samples, given a false positive rate  $p_{10}$ , is  $1 - (1 - p_{10})^T$  which increases rapidly as a function of  $p_{10}$  (demonstrated empirically by simulation, Table 3), and  $T$ . This was manifest in the results presented in Table 2, for which, with  $T = 11$ , there are many opportunities to observe a false positive.

A referee commented that our formulation of the model for misclassification suggests that false positive detections can only occur at sites where the species is not present, yet it seems plausible that they can also occur where the species is present. We did not explicitly address the possibility that a correct classification occurred by mistake because it does not affect the statistical inference with regard to site occupancy. In this case, the site is not misclassified. Indeed, under the

model allowing for false positives,  $p_{11}$  is not a “correct” detection rate parameter but rather a correct site classification rate parameter. This phenomenon does however modify the definition of  $p_{11}$  (detection probability for occupied sites) to be equal to the “net” probability of detection. That is, suppose  $p_a$  is the probability of correctly detecting a species at an occupied site and  $p_b$  is the probability of incorrectly detecting a species (at either an occupied site or an unoccupied site), and suppose that the two types of detection events are independent, then the likelihood for the data is precisely equivalent to that given by Eq. (1), but with  $p_{11} = p_a + p_b - p_a p_b$  and  $p_{00} = p_b$ . Both parameters  $p_a$  and  $p_b$  can be estimated, but this is not necessary when interest is focused on  $\psi$ .

The model that arises in the presence of false positive errors is closely related to conventional models of heterogeneous detection in animal surveys (commonly referred to collectively as model  $M_h$ ). Mathematically, the misclassification model is identical to the finite mixture of Norris and Pollock (1996) used to estimate the size of a closed population. The difference is that, here,  $N$  (the number of sample sites) is known. Similar models have been proposed for site occupancy models with heterogeneity (Royle 2005), and their form is slightly different. In the present case, the data are binomial counts with occupied sites having one value of  $p$  and unoccupied sites having another. In the presence of false positive errors, under a finite mixture model of order  $k$  for  $p$ , the resulting likelihood is precisely a finite mixture of order  $k + 1$ . For example, if  $p$  for occupied sites is a finite mixture with two components, and false positives are possible, the resulting likelihood is a mixture of three binomial distributions. On the other hand, in the absence of false positives, when variation in  $p$  is described by a finite mixture of order  $k$ , the resulting likelihood is a zero-inflated finite mixture of order  $k$  (see Royle 2005). Of course, alternative models for heterogeneity are possible (e.g., Coull and Agresti 1999, Dorazio and Royle 2003). In light of recent work by Link (2003) (see also Pledger [2005] and Dorazio and Royle [2005]), we suggest tempering enthusiasm for building intricate models of detection probability and over-reliance on conventional model-selection procedures as a basis for inference in such problems.

In general, site occupancy models, including the generalization that we have proposed, inherit certain mathematical properties from their common representation as binomial mixture models. This has the important consequence that there can be multiple interpretations of certain models. To bring this issue into context, we take the general class of models indexed by three parameters, site occupancy ( $\psi$ ), detection probability ( $p_{11}$ ) and the false positive error rate ( $p_{10}$ ). It is convenient to reference specific models from within this general class using the short-hand notation  $(\psi, p_{11}, p_{10})$ . In this context, the null model, i.e., that of MacKenzie et al. (2002), is the reduced model  $(\psi, p_{11}, p_{10} = 0)$ . We note

that  $(\psi, p_{11}, p_{10} = 0)$  is equivalent in terms of likelihood to the model  $(1 - \psi, p_{11} = 0, p_{10})$ , merely due to the symmetry in the likelihood that results from the binomial mixture structure. These competing interpretations are avoided in MacKenzie et al. (2002) and applications of that model, by ruling out the case  $p_{10} > 0$  a priori. We have noted that this is not a necessary constraint, as the condition  $p_{11} > p_{10}$  yields an unambiguous interpretation of the more general model, and the constrained model is falsifiable under the more general model. Competing interpretations also arise under our generalization of site occupancy models. For example, suppose we have an ordinary site occupancy model with heterogeneous detection probabilities described by a finite mixture of order  $k$ . This model is a zero-inflated binomial mixture (see Royle 2006), with occupancy parameter  $\psi$ , a false positive error rate of identically zero, and several parameters that describe the binomial mixture. Denote this model by  $(\psi, p_{11} \sim \text{FM}(k), p_{10} = 0)$ . The model generalized to allow for false positives is  $(\psi, p_{11} \sim \text{FM}(k), p_{10})$ . The more general model is precisely equivalent to the model  $(\psi = 1, p_{11} \sim \text{FM}(k + 1), p_{10} = 0)$  in the sense that any data will yield exactly the same likelihood.

How might we resolve the existence of ambiguous interpretations of such models? We could adopt the conventional approach of ruling out pathological boundary cases such as  $\psi = 1$  in order to yield an unambiguous interpretation of the model. We might also choose to admit the ambiguity in our assessment, as there is no formal, objective, basis for neglecting the case  $\psi = 1$  or the case  $p_{10} = 0$  in the conventional site occupancy model. Our view is that one has to consider the viability of competing interpretations in light of understanding of the sampling problem at hand. For example, Ovenbirds have a distinctive call, and it is hard to imagine that a well-trained observer produces false positive detection. Thus, one interpretation of the results given in Table 2 is that  $\psi \approx 1$  and that  $p_{11}$  is heterogeneous and approximated by a two-point finite-mixture distribution. In the present case, false positive errors and heterogeneous detection probabilities (parameterized by a finite mixture) cannot be considered as competing data generating mechanisms because one cannot distinguish between them from data. That is, no statistical analysis can yield information in support of one of the mechanisms over the other and so the model must be described a priori. The finite mixture as used in this paper has a direct mechanistic interpretation—it is not one of many possible candidate models of misclassification, but rather it is the model that results from the introduction of random false positive errors into the zero-inflated binomial sampling model underlying the MacKenzie et al. (2002) model. Conversely, in the classical heterogeneity models (e.g., Norris and Pollock 1996) the finite mixture is but one of many models of heterogeneous detection probabilities, and its application is typically as a curve-fitting tool employed to

improve model fit. We thus believe that its use to describe heterogeneous detection probabilities in site occupancy models should be avoided. We note that continuous heterogeneity models are identifiable in the presence of misclassification, the resulting likelihood being a mixture of the continuous density and a point mass at some nonzero value, i.e.,  $p_{10}$ .

#### LITERATURE CITED

- Bailey, L. L., T. R. Simons, and K. H. Pollock. 2004. Estimating site occupancy and species detection probability parameters for terrestrial salamanders. *Ecological Applications* **14**:692–702.
- Ball, L. C., P. F. Doherty, Jr., and M. W. McDonald. 2005. Using estimates of patch occupancy to evaluate a habitat model. *Journal of Wildlife Management* **69**(3):894–904.
- Bayley, P. B., and J. T. Peterson. 2001. An approach to estimate probability of presence and richness of fish species. *Transactions of the American Fisheries Society* **130**:620–633.
- Burnham, K. H., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer, New York, New York, USA.
- Coull, B. A., and A. Agresti. 1999. The use of mixed logit models to reflect heterogeneity in capture–recapture studies. *Biometrics* **55**:294–301.
- Dorazio, R. M., and J. A. Royle. 2003. Mixture models for estimating the size of a closed population when capture rates vary among individuals. *Biometrics* **59**:351–364.
- Dorazio, R. M., and J. A. Royle. 2005. Rejoinder to “The performance of mixture models in heterogeneous closed population capture–capture.” *Biometrics* **61**:874–876.
- Gu, W., and R. K. Swihart. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. *Biological Conservation* **116**:195–203.
- Hall, R. J., and C. A. Langtimm. 2001. The U. S. National Amphibian Research and Monitoring Initiative and the role of protected areas. *George Wright Forum* **18**:14–25.
- Link, W. A. 2003. Nonidentifiability of population size from capture–recapture data with heterogeneous detection probabilities. *Biometrics* **59**:1123–1130.
- MacKenzie, D. I., and L. L. Bailey. 2004. Assessing the fit of site occupancy models. *Journal of Agricultural, Biological and Ecological Statistics* **9**(3):300–318.
- MacKenzie, D. I., L. L. Bailey, and J. D. Nichols. 2005. Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology* **73**:546–555.
- MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. D. Franklin. 2003. Estimating site occupancy, colonization and local extinction when a species is detected imperfectly. *Ecology* **84**:2200–2207.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than 1. *Ecology* **83**:2248–2255.
- Nichols, J. D., and K. U. Karanth. 2002. Statistical concepts; assessing spatial distribution. Pages 29–38 in K. U. Karanth and J. D. Nichols, editors. *Monitoring tigers and their prey*. Centre for Wildlife Studies, Bangalore, India.
- Norris, J. L., III, and K. H. Pollock. 1996. Nonparametric MLE under two closed capture–recapture models with heterogeneity. *Biometrics* **52**:639–649.
- Pledger, S. 2005. The performance of mixture models in heterogeneous closed population capture–recapture. *Biometrics* **61**:868–876.
- R Development Core Team. 2005. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

- Royle, J. A. 2004a. N-mixture models for estimating population size from spatially replicated counts. *Biometrics* **60**:108–115.
- Royle, J. A. 2004b. Modeling abundance index data from anuran calling surveys. *Conservation Biology* **18**:1378–1385.
- Royle, J. A. 2006. Site occupancy models with heterogeneous detection probabilities. *Biometrics*, *in press*.
- Royle, J. A., and W. A. Link. 2005. A general class of multinomial mixture models for anuran calling survey data. *Ecology* **86**:2505–2512.
- Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. *Ecology* **84**:777–790.
- Stanley, T. R., and J. A. Royle. 2005. Estimating animal abundance using passive detection indices. *Journal of Wildlife Management* **69**(3):874–883.
- Swihart, R. K., T. C. Atwood, J. R. Goheen, D. M. Scheiman, K. E. Munroe, and T. M. Gehring. 2003. Patch occupancy of North American mammals: is patchiness in the eye of the beholder? *Journal of Biogeography* **30**:1259–1279.
- Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, and H. P. Possingham. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecological Applications* **13**:1790–1801.
- Weir, L. A., J. A. Royle, P. Nanjappa, and R. E. Jung. 2005. Modeling anuran detection and site occupancy on North American Amphibian Monitoring Program (NAAMP) routes in Maryland. *Journal of Herpetology* **39**:627–639.
- Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego, California, USA.

#### SUPPLEMENT

Data and R code for analyses summarized in Table 1 (*Ecological Archives* E087-049-S1).